(19) Europäisch s Patentamt
Eur pean Pat nt Offic
Offi européen des brev ts

(11) EP 1 033 405 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
06.09.2000 Bulletin 2000/36

(21) Application number: **00301439.6**

(22) Date of filing: **25.02.2000**

(51) Int Cl.[7]: **C12N 15/29**, C12N 15/82, C07K 14/415, C12Q 1/68, A01H 5/00

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **25.02.1999 US 121825 P**
27.07.1999 US 145918 P
28.07.1999 US 145951 P
02.08.1999 US 146388 P
02.08.1999 US 146389 P
02.08.1999 US 146386 P
03.08.1999 US 147038 P
04.08.1999 US 147302 P
04.08.1999 US 147204 P
More priorities on the following pages

(83) **Declaration under Rule 28(4) EPC (expert solution)**

(71) Applicant: **Ceres Incorporated**
Malibu, CA 90265 (US)

(72) Inventors:
- **Alexandrov, Nickolai**
  Thousand Oaks, CA 91320 (US)
- **Brover, Vyacheslav**
  Calabasas, CA 91302 (US)
- **Chen, Xianfeng**
  Los Angeles, CA 90025 (US)
- **Subramanian, Gopalakrishnan**
  Moorpark, CA 93021 (US)
- **Troukhan, Maxim E.**
  South Pasadena, CA 91030 (US)
- **Zheng, Liansheng**
  Creve Coeur, MO 63141 (US)
- **Dumas, J.**
  , (US)

(74) Representative:
Bannerman, David Gardner et al
Withers & Rogers,
Goldings House,
2 Hays Lane
London SE1 2HW (GB)

Remarks:
THE COMPLETE DOCUMENT INCLUDING REFERENCE TABLES AND THE SEQUENCE LISTING IS AVAILABLE ON CD-ROM FROM THE EUROPEAN PATENT OFFICE, VIENNA SUB-OFFICE.

(54) Sequence-determined DNA fragments and corresponding polypeptides encoded thereby

(57) The present invention provides DNA molecules that constitute fragments of the genome of a plant, and polypeptides encoded thereby. The DNA molecules are useful for specifying a gene product in cells, either as a promoter or as a protein coding sequence or as an UTR or as a 3' termination sequence, and are also useful in controlling the behavior of a gene in the chromosome, in controlling the expression of a gene or as tools for genetic mapping, recognizing or isolating identical or r - lated DNA fragments, or identification of a particular individual organism, or for clustering of a group of organisms with a common trait.

[0]Arabidopsis DNA is used in the present experiment, but the procedure is a general one.

EP 1 033 405 A2

(30) Priority: Continued from first pag
05.08.1999 US 147260 P
05.08.1999 US 147192 P
06.08.1999 US 147303 P
06.08.1999 US 147416 P
09.08.1999 US 147493 P
09.08.1999 US 147935 P
10.08.1999 US 148171 P
11.08.1999 US 148319 P
12.08.1999 US 148341 P
13.08.1999 US 148565 P
13.08.1999 US 148684 P
16.08.1999 US 149368 P
17.08.1999 US 149175 P
18.08.1999 US 149426 P
20.08.1999 US 149722 P
20.08.1999 US 149929 P
20.08.1999 US 149723 P
23.08.1999 US 149902 P
23.08.1999 US 149930 P
25.08.1999 US 150566 P
26.08.1999 US 150884 P
27.08.1999 US 151065 P
27.08.1999 US 151066 P
27.08.1999 US 151080 P
30.08.1999 US 151303 P
31.08.1999 US 151438 P
01.09.1999 US 151930 P
07.09.1999 US 152363 P
10.09.1999 US 153070 P
13.09.1999 US 153758 P
15.09.1999 US 154018 P
16.09.1999 US 154039 P
20.09.1999 US 154779 P
22.09.1999 US 155139 P
23.09.1999 US 155486 P
24.09.1999 US 155659 P
28.09.1999 US 156458 P
29.09.1999 US 156596 P
04.10.1999 US 157117 P
05.10.1999 US 157753 P
06.10.1999 US 157865 P
07.10.1999 US 158029 P
08.10.1999 US 158232 P
12.10.1999 US 158369 P
13.10.1999 US 159294 P
13.10.1999 US 159295 P
13.10.1999 US 159293 P
14.10.1999 US 159638 P
14.10.1999 US 159637 P
14.10.1999 US 159329 P
14.10.1999 US 159331 P
14.10.1999 US 159330 P
18.10.1999 US 159584 P
21.10.1999 US 160815 P
21.10.1999 US 160767 P
21.10.1999 US 160768 P
21.10.1999 US 160741 P
21.10.1999 US 160770 P
21.10.1999 US 160814 P
22.10.1999 US 160981 P
22.10.1999 US 160980 P
22.10.1999 US 160989 P
25.10.1999 US 161405 P
25.10.1999 US 161404 P
25.10.1999 US 161406 P
26.10.1999 US 161361 P
26.10.1999 US 161360 P
26.10.1999 US 161359 P
28.10.1999 US 161920 P
28.10.1999 US 161992 P
28.10.1999 US 161993 P
29.10.1999 US 162143 P
29.10.1999 US 162142 P
29.10.1999 US 162228 P
01.11.1999 US 162895 P
01.11.1999 US 162891 P
01.11.1999 US 162894 P
02.11.1999 US 163093 P
02.11.1999 US 163092 P
02.11.1999 US 163091 P
03.11.1999 US 163249 P
03.11.1999 US 163248 P
03.11.1999 US 163281 P
04.11.1999 US 163380 P
04.11.1999 US 163381 P
04.11.1999 US 163379 P
08.11.1999 US 164151 P
08.11.1999 US 164150 P
08.11.1999 US 164146 P
09.11.1999 US 164260 P
09.11.1999 US 164259 P
10.11.1999 US 164548 P
10.11.1999 US 164317 P
10.11.1999 US 164321 P
10.11.1999 US 164318 P
10.11.1999 US 164544 P
10.11.1999 US 164545 P
10.11.1999 US 164319 P
12.11.1999 US 164870 P
12.11.1999 US 164959 P
12.11.1999 US 164962 P
12.11.1999 US 164960 P
12.11.1999 US 164871 P
12.11.1999 US 164961 P
15.11.1999 US 164927 P
15.11.1999 US 164929 P
15.11.1999 US 164926 P
16.11.1999 US 165669 P
16.11.1999 US 165671 P
16.11.1999 US 165661 P
17.11.1999 US 165919 P
17.11.1999 US 165918 P
17.11.1999 US 165911 P
18.11.1999 US 166158 P
18.11.1999 US 166157 P

(Cont. next pag)

18.11.1999 US 166173 P
19.11.1999 US 166412 P
19.11.1999 US 166419 P
19.11.1999 US 166411 P
22.11.1999 US 166733 P
22.11.1999 US 166750 P
23.11.1999 US 167362 P
24.11.1999 US 167382 P
24.11.1999 US 167233 P
24.11.1999 US 167234 P
24.11.1999 US 167235 P
30.11.1999 US 167904 P
30.11.1999 US 167908 P
30.11.1999 US 167902 P
01.12.1999 US 168232 P
01.12.1999 US 168233 P
01.12.1999 US 168231 P
02.12.1999 US 168546 P
02.12.1999 US 168549 P
02.12.1999 US 168548 P
03.12.1999 US 168673 P
03.12.1999 US 168675 P
03.12.1999 US 168674 P
07.12.1999 US 169278 P
07.12.1999 US 169302 P
07.12.1999 US 169298 P
08.12.1999 US 169692 P
08.12.1999 US 169691 P
16.12.1999 US 171107 P
16.12.1999 US 171098 P
16.12.1999 US 171114 P
19.01.2000 US 176866 P
19.01.2000 US 176867 P
19.01.2000 US 176910 P
26.01.2000 US 178166 P
27.01.2000 US 178547 P
27.01.2000 US 177666 P
27.01.2000 US 178546 P
27.01.2000 US 178544 P
27.01.2000 US 178545 P
28.01.2000 US 178755 P
28.01.2000 US 178754 P
01.02.2000 US 179395 P
01.02.2000 US 179388 P
03.02.2000 US 180039 P
03.02.2000 US 180139 P
04.02.2000 US 180207 P
04.02.2000 US 180206 P
07.02.2000 US 180695 P
07.02.2000 US 180696 P
09.02.2000 US 181228 P
09.02.2000 US 181214 P
10.02.2000 US 181476 P
10.02.2000 US 181551 P
15.02.2000 US 182477 P
15.02.2000 US 182516 P
15.02.2000 US 182512 P
15.02.2000 US 182478 P
17.02.2000 US 183165 P
17.02.2000 US 183166 P
27.07.1999 US 145913 P
05.03.1999 US 123180 P
09.03.1999 US 123548 P
23.03.1999 US 125788 P
25.03.1999 US 126264 P
29.03.1999 US 126785 P
01.04.1999 US 127462 P
06.04.1999 US 128234 P
08.04.1999 US 128714 P
16.04.1999 US 129845 P
19.04.1999 US 130077 P
21.04.1999 US 130449 P
23.04.1999 US 130891 P
23.04.1999 US 130510 P
28.04.1999 US 131449 P
30.04.1999 US 132407 P
30.04.1999 US 132048 P
04.05.1999 US 132484 P
05.05.1999 US 132485 P
06.05.1999 US 132487 P
06.05.1999 US 132486 P
07.05.1999 US 132863 P
11.05.1999 US 134256 P
14.05.1999 US 134221 P
14.05.1999 US 134218 P
14.05.1999 US 134370 P
14.05.1999 US 134219 P
18.05.1999 US 134768 P
19.05.1999 US 134941 P
20.05.1999 US 135124 P
21.05.1999 US 135353 P
24.05.1999 US 135629 P
25.05.1999 US 136021 P
27.05.1999 US 136392 P
28.05.1999 US 136782 P
01.06.1999 US 137222 P
03.06.1999 US 137528 P
04.06.1999 US 137502 P
07.06.1999 US 137724 P
08.06.1999 US 138094 P
10.06.1999 US 138540 P
10.06.1999 US 138847 P
14.06.1999 US 139119 P
16.06.1999 US 139452 P
16.06.1999 US 139453 P
17.06.1999 US 139492 P
18.06.1999 US 139461 P
18.06.1999 US 139750 P
18.06.1999 US 139463 P
18.06.1999 US 139457 P
18.06.1999 US 139459 P
18.06.1999 US 139462 P
18.06.1999 US 139455 P
18.06.1999 US 139458 P
18.06.1999 US 139454 P
18.06.1999 US 139456 P

(Cont. next pag)

18.06.1999 US 139460 P
18.06.1999 US 139763 P
21.06.1999 US 139817 P
22.06.1999 US 139899 P
23.06.1999 US 140354 P
23.06.1999 US 140353 P
24.06.1999 US 140695 P
28.06.1999 US 140823 P
29.06.1999 US 140991 P
30.06.1999 US 141287 P
01.07.1999 US 142154 P
01.07.1999 US 141842 P
02.07.1999 US 142055 P
06.07.1999 US 142390 P
08.07.1999 US 142803 P
09.07.1999 US 142920 P
12.07.1999 US 142977 P
13.07.1999 US 143542 P
14.07.1999 US 143624 P
15.07.1999 US 144005 P
16.07.1999 US 144085 P
16.07.1999 US 144086 P
19.07.1999 US 144333 P
19.07.1999 US 144335 P
19.07.1999 US 144325 P
19.07.1999 US 144334 P
19.07.1999 US 144332 P
19.07.1999 US 144331 P
20.07.1999 US 144884 P
20.07.1999 US 144352 P
20.07.1999 US 144632 P
21.07.1999 US 144814 P
21.07.1999 US 145086 P
21.07.1999 US 145088 P
22.07.1999 US 145192 P
22.07.1999 US 145085 P
22.07.1999 US 145089 P
22.07.1999 US 145087 P
23.07.1999 US 145145 P
23.07.1999 US 145224 P
23.07.1999 US 145218 P
26.07.1999 US 145276 P
27.07.1999 US 145919 P

[2347] Th suspension culture cells are transformed with exogenous DNA as described by Z. Chen et al. *Plant Mol. Bio.* 36:163 (1998). Briefly, 4-days post-subculture cells are incubated with cell wall digestion solution containing 0.4 M sorbitol, 2% driselase, 5mM MES (2-[N-Morpholino] ethanesulfonic acid) pH 5.0 for 5 hours. The digest d cells are pelleted gently at 60 xg for 5 min. and washed twice in W5 solution containing 154 mM NaCl, 5 mM KCl, 125 mM $CaCl_2$ and 5mM glucose, pH 6.0. Th protoplasts are suspend d in MC solution containing 5 mM MES, 20 mM $CaCl_2$, 0.5 M mannitol, pH 5.7 and the protoplast density is adjusted to about 4 x 10^6 protoplasts per ml.

[2348] 15-60 µg of plasmid DNA is mixed with 0.9 ml of protoplasts. The resulting suspension is mixed with 40% polyethylene glycol (MW 8000, PEG 8000), by gentle inversion a few times at room temperature for 5 to 25 min. Protoplast culture medium known in the art is added into the PEG-DNA-protoplast mixture. Protoplasts are incubat d in the culture medium for 24 hour to 5 days and cell extracts can be used for assay of transient expression of the introduced gene. Alternatively, transformed cells can be used to produce transgenic callus, which in turn can be used to produce transgenic plants, by methods known in the art. See, for example, Nomura and Komamine, *Pit. Phys.* 79: 988-991 (1985), *Identification and Isolation of Single Cells that Produce Somatic Embryos in Carrot Suspension Cultures.*

[2349] The invention being thus described, it will be apparent to one of ordinary skill in the art that various modifications of the materials and methods for practicing the invention can be made. Such modifications are to be considered within the scope of the invention as defined by the following claims.

[2350] Each of the references from the patent and periodical literature cited herein is hereby expressly incorporated in its entirety by such citation.

Claims

1. An isolated nucleic acid molecule comprising a nucleic acid having a nucleotide sequence which encodes an amino acid sequence exhibiting at least 40% sequence identity to an amino acid sequence encoded by

   (a) a nucleotide sequence described in REF and/or SEQ Table 1 or 2 or a fragment thereof; or
   (b) a complement of a nucleotide sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof.

2. An isolated nucleic acid molecule comprising a nucleic acid having a nucleotide sequence which exhibits at least 65% sequence identity to

   (a) a nucleotide sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof; or
   (b) a complement of a nucleotide sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof.

3. An isolated nucleic acid molecule comprising a nucleic acid having a nucleotide sequence which exhibits at least 65% sequence identity to a gene comprising

   (a) a nucleotide sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof; or
   (b) a complement of a nucleotide sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof.

4. An isolated nucleic acid molecule which is the reverse of the isolated nucleotide sequence according to any one of claims 1-3, such that the reverse nucleotide sequence has a sequence order which is the reverse of the sequence order of said isolated nucleotide sequence according to any one of claims 1-3.

5. An isolated nucleic acid molecule comprising a nucleic acid capable of hybridizing to a nucleic acid having a sequence selected from the group consisting of:

   (a) a nucleotide sequence which is shown in REF and/or SEQ Table 1 or 2; and
   (b) a nucleotide sequence which is complementary to a nucleotide sequence shown in REF and/or SEQ Table 1 or 2;

   under conditions that permit formation of a nucleic acid duplex at a temperature from about 40°C and 48°C below the melting temperature of the nucleic acid duplex.

6. The nucleic acid molecule according to any one of claims 1-5, wherein said nucleic acid comprises an open reading frame.

7. The isolated nucleic acid molecule of any one of claims 1-5, wherein said nucleic acid is capable of functioning as a promoter, a 3' end termination sequence, an untranslated region (UTR), or as a regulatory sequence.

8. The isolated nucleic acid molecule of claim 7, wherein said nucleic acid is a promoter and comprises a sequence selected from the group consisting of a TATA box sequence, a CAAT box sequence, a motif of GCAATCG or any transcriptoin-factor binding sequence, and any combination thereof.

9. The isolated nucleic acid molecule of claim 7, wherein the nucleic acid sequence is a regulatory sequence which is capable of promoting seed-specific expression, embryo-specific expression, ovule-specific expression, tapetum-specific expression or root-specific expression of a sequence or any combination thereof.

10. A vector construct comprising a nucleic acid molecule according to any one of claims 1-9, wherein said nucleic acid molecule is heterologous to any element in said vector construct.

11. A vector construct according to claim 10 comprising:

(a) a first nucleic acid having a regulatory sequence capable of causing transcription and/or translation; and
(b) a second nucleic acid having the sequence of said isolated nucleic acid molecule according to any one of claims 1-4;

wherein said first and second nucleic acids are operably linked and wherein said second nucleic acid is heterologous to any element in said vector construct.

12. The vector construct according to claim 11, wherein said first nucleic acid is native to said second nucleic acid.

13. The vector construct according to claim 11, wherein said first nucleic acid is heterologous to said second nucleic acid.

14. A vector construct according to claim 10 comprising:

(c) a first nucleic acid having having the sequence of said isolated nucleic acid molecule according to claim 7; and
(d) a second nucleic acid;

wherein said first and second nucleic acids are operably linked and wherein said first nucleic acid is heterologous to any element in said vector construct.

15. The vector construct according to claim 14, wherein said first nucleic acid is native to said second nucleic acid.

16. The vector construct according to claim 14, wherein said first nucleic acid is heterologous to said second nucleic acid.

17. A host cell comprising an isolated nucleic acid molecule according to any one of claims 1-4, wherein said nucleic acid molecule is flanked by exogenous sequence.

18. A host cell comprising a vector construct of any one of claims 10-16.

19. An isolated polypeptide comprising an amino acid sequence

(a) exhibiting at least 40% sequence identity of an amino acid sequence encoded by a sequence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof; and
(b) capable of exhibiting at least one of the biological activities of the polypeptide encoded by said nucleotide seqence shown in REF and/or SEQ Table 1 or 2 or a fragment thereof.

20. The isolated polypeptide of claim 19, wherein said amino acid sequence exhibits at least 75% sequence identity to an amino acid sequence encoded by a sequence shown in SEQ Table 1 or 2 or a fragment thereof.

21. The isolated polypeptide of claim 19, wherein said amino acid sequence exhibits at least 85% sequenc identity

to an amino acid sequence encoded by a sequence shown in SEQ Table 1 or 2 or a fragment thereof.

22. The isolated polypeptide of claim 19, wherein said amino acid sequence exhibits at least 90% sequence identity to an amino acid sequence encoded by a sequence shown in SEQ Table 1 or 2 or a fragment ther of.

23. An antibody capable of binding the isolated polypeptide of any one of claims 19-22.

24. A method of introducing an isolated nucleic acid into a host cell comprising:

(a) providing an isolated nucleic acid molecule according to any one of claims 1-4; and
(b) contacting said isolated nucleic with said host cell under conditions that permit insertion of said nucleic acid into said host cell.

25. A method of transforming a host cell which comprises contacting a host cell with a vector construct according to any one of claims 10-16.

26. A method of modulating transcription and/or translation of a nucleic acid in a host cell comprising:

(a) providing the host cell of claim 24 or 25; and
(b) culturing said host cell under conditions that permit transcription or translation.

27. A method for detecting a nucleic acid in a sample which comprises:

(a) providing an isolated nucleic acid molecule according to any one of claims 1-5;
(b) contacting said isolated nucleic acid molecule with a sample under conditions which permit a comparison of the sequence of said isolated nucleic acid molecule with the sequence of DNA in said sample; and
(c) analyzing the result of said comparison.

28. The method according to claim 27, wherein said isolated nucleic acid molecule and said sample are contacted under conditions which permit the formation of a duplex between complementary nucleic acid sequences.

29. A plant or cell of a plant which comprises a nucleic acid molecule according to any one of claims 1-4 which is exogenous to said plant or plant cell.

30. A plant or cell of a plant which comprises a nucleic acid molecule according to any one of claims 1-4, wherein said nucleic acid molecule is heterologous to said plant or said cell of a plant.

31. A plant or cell of a plant which has been transformed with a nucleic acid molecule according to any one of claims 1-4.

32. A plant of cell of a plant which comprises a vector construct according to any one of claims 10-16.

33. A plant of cell of a plant which has been transformed with a vector construct according to any one of claims 10-16.

34. A plant which has been regenerated from a plant cell according to any one of claims 29-33.

- Align. NO 19629
- gi No 3668069
- Desp. : (U28007) Pto kinase interactor 1 [Lycopersicon esculentum]
- % Idnt. : 73.2
- Align. Len.: 327
- Loc. SEQ ID NO 28400: 1 -> 204 aa.

Max Len. Seq. :
rel to:
Clone IDs:
34671
(Ac) cDNA SEQ
- Pat. Appln. SEQ ID NO: 28401
- Ceres SEQ ID NO: 1571425

PolyP SEQ
- Pat. Appln. SEQ ID NO 28402
- Ceres SEQ ID NO 1571426
- Loc. SEQ ID NO 28401: @ 3 nt.

(C) Pred. PP Nom. & Annot.
(Dp) Rel. AA SEQ
- Align. NO 19630
- gi No 3047124
- Desp. : (AF058919) No definition line found [Arabidopsis thaliana]
- % Idnt. : 100
- Align. Len.: 68
- Loc. SEQ ID NO 28402: 173 -> 240 aa.

PolyP SEQ
- Pat. Appln. SEQ ID NO 28403
- Ceres SEQ ID NO 1571427
- Loc. SEQ ID NO 28401: @ 27 nt.

(C) Pred. PP Nom. & Annot.
(Dp) Rel. AA SEQ
- Align. NO 19631
- gi No 3047124
- Desp. : (AF058919) No definition line found [Arabidopsis thaliana]
- % Idnt. : 100
- Align. Len.: 68
- Loc. SEQ ID NO 28403: 165 -> 232 aa.

PolyP SEQ
- Pat. Appln. SEQ ID NO 28404
- Ceres SEQ ID NO 1571428
- Loc. SEQ ID NO 28401: @ 345 nt.

(C) Pred. PP Nom. & Annot.
(Dp) Rel. AA SEQ
- Align. NO 19632
- gi No 3047124
- Desp. : (AF058919) No definition line found [Arabidopsis thaliana]
- % Idnt. : 100
- Align. Len.: 68
- Loc. SEQ ID NO 28404: 59 -> 126 aa.

Max Len. Seq. :
rel to:
Clone IDs:

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..217 / Ceres Seq. ID 1571424

<400> 28400

```
Met Gly Ser Leu His Asp Ile Leu His Gly Arg Lys Gly Val Gln Gly
1               5                  10                  15
Ala Gln Pro Gly Pro Thr Leu Asp Trp Ile Gln Arg Val Arg Ile Ala
            20                  25                  30
Val Asp Ala Ala Arg Gly Leu Glu Tyr Leu His Glu Lys Val Gln Pro
        35                  40                  45
Ala Val Ile His Arg Asp Ile Arg Ser Ser Asn Val Leu Leu Phe Glu
    50                  55                  60
Asp Phe Lys Ala Lys Ile Ala Asp Phe Asn Leu Ser Asn Gln Ser Pro
65                  70                  75                  80
Asp Met Ala Ala Arg Leu His Ser Thr Arg Val Leu Gly Xaa Phe Gly
                85                  90                  95
Tyr His Ala Pro Glu Tyr Ala Met Thr Gly Gln Leu Thr Gln Lys Ser
            100                 105                 110
Asp Val Tyr Ser Phe Gly Val Val Leu Leu Glu Leu Leu Thr Gly Arg
        115                 120                 125
Lys Pro Val Asp His Thr Met Pro Arg Gly Gln Gln Ser Leu Val Thr
    130                 135                 140
Trp Ala Thr Pro Arg Leu Ser Glu Asp Lys Val Lys Gln Cys Val Asp
145                 150                 155                 160
Pro Lys Leu Lys Gly Glu Tyr Pro Pro Lys Ala Val Ala Lys Leu Ala
                165                 170                 175
Ala Val Ala Ala Leu Cys Val Gln Tyr Glu Ser Glu Phe Arg Pro Asn
            180                 185                 190
Met Ser Ile Val Val Lys Ala Leu Gln Pro Leu Leu Arg Ser Ser Thr
        195                 200                 205
Ala Ala Ala Val Pro Val Gln Glu Ala
    210                 215
```

<210> 28401
<211> 1500
<212> DNA
<213> Arabidopsis thaliana

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..1500 / Ceres Seq. ID 1571425

<400> 28401

```
aacgtttcga cagtctctac accgtcatga acactgaatc agttgtcgag ttccttggga  60
atgtgacctt gttgcagcgg ttacctagtt cctctctgaa gagaatctcc gaagtcgttg  120
tcttcaaagg ttatgacaga ggtgattatg tggttcgtga aaatcaaaat gtggatggag  180
tttatttttct cttgcaagga caggctcagg ttctgagatc agccgaagag gaaaactatc  240
aagagttccc tttgaaacga tatgatttct tcggccatgg tattttcggg gatgtttact  300
cagcagatgt tgttgctgtg acagagctta cctgcttgct gttgatgtct gatcatcgtg  360
ctttacttga aataaagtca gtctcggatt cagataagga acgctgtctt gtggaagaca  420
tactatatct agaaccatta gatttgaatg tataccgggg gttcacccca cctaatgctc  480
caacctatgg aaaggtttat ggagggcaat tagttggaca ggcacttgcc gcagcatcaa  540
aaactgttga aactatgaag atagtccata attttcattg ctatttcctt cttgttggag  600
atataaatat tcccatcata tatgatgtta accgcttacg tgacggcaac aactttgcca  660
```

```
ccagaagtgt agatgctaga cagaaaggaa aaactatatt caccttgttc gcgtcatttc  720
agaaaaagca acaaggtttt attcaccagg agtcgaccat gcctcataca ccagctcctg  780
aaacgcttct accaagggag gagatgcttg aacggcttgt tactgagcct ctgctaccta  840
gggattaccg aaaccaagtt gcaactgaaa ttagtgttcc attccctata gatattcgat  900
tttgtgagcc aaatcgttcc actaaacaga ataagtctcc tccaagacta aaatattggt  960
ttagagcaaa gggaaaactt tctgatgatg atcaagcttt gcacagatgt gtggttgcat 1020
ttgcttccga tttgatattc gccactatca gtttaaaccc tcaccggaga gagggcatga 1080
gtgtagctgc tcttagcctg gaccactcga tgtggttcca ccgacctgta agagcagatg 1140
attggttgtt gtttgtgtga gtccaactgc gacctaaagc cgcggttttg caactggcaa 1200
aatgttcaac agaaagggag agctggtggt atcattgacg caagaagctg tgttaagaga 1260
agctgtgact attaagccat cCttcggggc caagctatga agccatagga ttttgatagt 1320
gagagaattg ctgcatctgt tactcctcac ggtcacattc caaagagtcg tcacttatac 1380
tacatttgca tatgtttttc gatccacaat tattattttc ccctctaaaa gggtctacac 1440
atatgtttgt ttgtaaccac gataatgttt caacagcaat gaaaaagcaa acagtggttc 1500
```

<210> 28402
<211> 385
<212> PRT
<213> Arabidopsis thaliana

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..385 / Ceres Seq. ID 1571426

<400> 28402

```
Arg Phe Asp Ser Leu Tyr Thr Val Met Asn Thr Glu Ser Val Val Glu
1               5                  10                  15
Phe Leu Gly Asn Val Thr Leu Leu Gln Arg Leu Pro Ser Ser Ser Leu
            20                  25                  30
Lys Arg Ile Ser Glu Val Val Val Phe Lys Gly Tyr Asp Arg Gly Asp
        35                  40                  45
Tyr Val Val Arg Glu Asn Gln Asn Val Asp Gly Val Tyr Phe Leu Leu
    50                  55                  60
Gln Gly Gln Ala Gln Val Leu Arg Ser Ala Glu Glu Glu Asn Tyr Gln
65                  70                  75                  80
Glu Phe Pro Leu Lys Arg Tyr Asp Phe Phe Gly His Gly Ile Phe Gly
                85                  90                  95
Asp Val Tyr Ser Ala Asp Val Val Ala Val Thr Glu Leu Thr Cys Leu
            100                 105                 110
Leu Leu Met Ser Asp His Arg Ala Leu Leu Glu Ile Lys Ser Val Ser
        115                 120                 125
Asp Ser Asp Lys Glu Arg Cys Leu Val Glu Asp Ile Leu Tyr Leu Glu
    130                 135                 140
Pro Leu Asp Leu Asn Val Tyr Arg Gly Phe Thr Pro Pro Asn Ala Pro
145                 150                 155                 160
Thr Tyr Gly Lys Val Tyr Gly Gly Gln Leu Val Gly Gln Ala Leu Ala
                165                 170                 175
Ala Ala Ser Lys Thr Val Glu Thr Met Lys Ile Val His Asn Phe His
            180                 185                 190
Cys Tyr Phe Leu Leu Val Gly Asp Ile Asn Ile Pro Ile Ile Tyr Asp
        195                 200                 205
Val Asn Arg Leu Arg Asp Gly Asn Asn Phe Ala Thr Arg Ser Val Asp
    210                 215                 220
Ala Arg Gln Lys Gly Lys Thr Ile Phe Thr Leu Phe Ala Ser Phe Gln
225                 230                 235                 240
Lys Lys Gln Gln Gly Phe Ile His Gln Glu Ser Thr Met Pro His Thr
                245                 250                 255
```

```
Pro Ala Pro Glu Thr Leu Leu Pro Arg Glu Glu Met Leu Glu Arg Leu
            260                 265                 270
Val Thr Glu Pro Leu Leu Pro Arg Asp Tyr Arg Asn Gln Val Ala Thr
        275                 280                 285
Glu Ile Ser Val Pro Phe Pro Ile Asp Ile Arg Phe Cys Glu Pro Asn
    290                 295                 300
Arg Ser Thr Lys Gln Asn Lys Ser Pro Pro Arg Leu Lys Tyr Trp Phe
305                 310                 315                 320
Arg Ala Lys Gly Lys Leu Ser Asp Asp Asp Gln Ala Leu His Arg Cys
                325                 330                 335
Val Val Ala Phe Ala Ser Asp Leu Ile Phe Ala Thr Ile Ser Leu Asn
            340                 345                 350
Pro His Arg Arg Glu Gly Met Ser Val Ala Ala Leu Ser Leu Asp His
        355                 360                 365
Ser Met Trp Phe His Arg Pro Val Arg Ala Asp Asp Trp Leu Leu Phe
    370                 375                 380
Val
385
```

<210> 28403
<211> 377
<212> PRT
<213> Arabidopsis thaliana

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..377 / Ceres Seq. ID 1571427

<400> 28403

```
Met Asn Thr Glu Ser Val Val Glu Phe Leu Gly Asn Val Thr Leu Leu
1               5                   10                  15
Gln Arg Leu Pro Ser Ser Ser Leu Lys Arg Ile Ser Glu Val Val Val
            20                  25                  30
Phe Lys Gly Tyr Asp Arg Gly Asp Tyr Val Val Arg Glu Asn Gln Asn
        35                  40                  45
Val Asp Gly Val Tyr Phe Leu Leu Gln Gly Gln Ala Gln Val Leu Arg
    50                  55                  60
Ser Ala Glu Glu Glu Asn Tyr Gln Glu Phe Pro Leu Lys Arg Tyr Asp
65                  70                  75                  80
Phe Phe Gly His Gly Ile Phe Gly Asp Val Tyr Ser Ala Asp Val Val
                85                  90                  95
Ala Val Thr Glu Leu Thr Cys Leu Leu Leu Met Ser Asp His Arg Ala
            100                 105                 110
Leu Leu Glu Ile Lys Ser Val Ser Asp Ser Asp Lys Glu Arg Cys Leu
        115                 120                 125
Val Glu Asp Ile Leu Tyr Leu Glu Pro Leu Asp Leu Asn Val Tyr Arg
    130                 135                 140
Gly Phe Thr Pro Pro Asn Ala Pro Thr Tyr Gly Lys Val Tyr Gly Gly
145                 150                 155                 160
Gln Leu Val Gly Gln Ala Leu Ala Ala Ala Ser Lys Thr Val Glu Thr
                165                 170                 175
Met Lys Ile Val His Asn Phe His Cys Tyr Phe Leu Leu Val Gly Asp
            180                 185                 190
Ile Asn Ile Pro Ile Ile Tyr Asp Val Asn Arg Leu Arg Asp Gly Asn
        195                 200                 205
Asn Phe Ala Thr Arg Ser Val Asp Ala Arg Gln Lys Gly Lys Thr Ile
    210                 215                 220
```

```
Phe Thr Leu Phe Ala Ser Phe Gln Lys Lys Gln Gln Gly Phe Ile His
225                 230                 235                 240
Gln Glu Ser Thr Met Pro His Thr Pro Ala Pro Glu Thr Leu Leu Pro
                245                 250                 255
Arg Glu Glu Met Leu Glu Arg Leu Val Thr Glu Pro Leu Leu Pro Arg
            260                 265                 270
Asp Tyr Arg Asn Gln Val Ala Thr Glu Ile Ser Val Pro Phe Pro Ile
        275                 280                 285
Asp Ile Arg Phe Cys Glu Pro Asn Arg Ser Thr Lys Gln Asn Lys Ser
    290                 295                 300
Pro Pro Arg Leu Lys Tyr Trp Phe Arg Ala Lys Gly Lys Leu Ser Asp
305                 310                 315                 320
Asp Asp Gln Ala Leu His Arg Cys Val Val Ala Phe Ala Ser Asp Leu
                325                 330                 335
Ile Phe Ala Thr Ile Ser Leu Asn Pro His Arg Arg Glu Gly Met Ser
            340                 345                 350
Val Ala Ala Leu Ser Leu Asp His Ser Met Trp Phe His Arg Pro Val
        355                 360                 365
Arg Ala Asp Asp Trp Leu Leu Phe Val
    370                 375


<210> 28404
<211> 271
<212> PRT
<213> Arabidopsis thaliana

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..271 / Ceres Seq. ID 1571428

<400> 28404
Met Ser Asp His Arg Ala Leu Leu Glu Ile Lys Ser Val Ser Asp Ser
1               5                   10                  15
Asp Lys Glu Arg Cys Leu Val Glu Asp Ile Leu Tyr Leu Glu Pro Leu
            20                  25                  30
Asp Leu Asn Val Tyr Arg Gly Phe Thr Pro Pro Asn Ala Pro Thr Tyr
        35                  40                  45
Gly Lys Val Tyr Gly Gly Gln Leu Val Gly Gln Ala Leu Ala Ala Ala
    50                  55                  60
Ser Lys Thr Val Glu Thr Met Lys Ile Val His Asn Phe His Cys Tyr
65                  70                  75                  80
Phe Leu Leu Val Gly Asp Ile Asn Ile Pro Ile Ile Tyr Asp Val Asn
                85                  90                  95
Arg Leu Arg Asp Gly Asn Asn Phe Ala Thr Arg Ser Val Asp Ala Arg
            100                 105                 110
Gln Lys Gly Lys Thr Ile Phe Thr Leu Phe Ala Ser Phe Gln Lys Lys
        115                 120                 125
Gln Gln Gly Phe Ile His Gln Glu Ser Thr Met Pro His Thr Pro Ala
    130                 135                 140
Pro Glu Thr Leu Leu Pro Arg Glu Glu Met Leu Glu Arg Leu Val Thr
145                 150                 155                 160
Glu Pro Leu Leu Pro Arg Asp Tyr Arg Asn Gln Val Ala Thr Glu Ile
                165                 170                 175
Ser Val Pro Phe Pro Ile Asp Ile Arg Phe Cys Glu Pro Asn Arg Ser
            180                 185                 190
Thr Lys Gln Asn Lys Ser Pro Pro Arg Leu Lys Tyr Trp Phe Arg Ala
        195                 200                 205
```

```
Lys Gly Lys Leu Ser Asp Asp Asp Gln Ala Leu His Arg Cys Val Val
    210                 215                 220
Ala Phe Ala Ser Asp Leu Ile Phe Ala Thr Ile Ser Leu Asn Pro His
225                 230                 235                 240
Arg Arg Glu Gly Met Ser Val Ala Ala Leu Ser Leu Asp His Ser Met
                245                 250                 255
Trp Phe His Arg Pro Val Arg Ala Asp Asp Trp Leu Leu Phe Val
            260                 265                 270
```

<210> 28405
<211> 1974
<212> DNA
<213> Arabidopsis thaliana

<220>
<223> any n or Xaa = unknown

<220>
<223> Location 1..1974 / Ceres Seq. ID 1571432

<400> 28405

```
aaaaaagaat acacagacca gtcttcgtgt gctcgtaatt ggcgattctc aaattcgatc  60
tccgtacaaa caatggccga aacctcgaag caagtcaacg gcgacgatgc ccaggatctt  120
cactcgcttc tttcttctcc ggcgagggat ttcctcgttc gtaatgacgg cgaacaggtg  180
aaagttgaca gcttgttagg gaagaagatt ggattgtatt tttcagctgc ttggtgtgga  240
ccgtgtcagc ggtttactcc acagctggtg gaagtctaca acgagctctc ttcgaaagtt  300
ggttttgaga ttgtgtttgt gtcaggtgat gaggatgaag agtcatttgg agattatttc  360
agtaagatgc cctggctcgc tgttccgttt actgattcgg aaacccgtga ccgtttggat  420
gagttgttta aggttagggg aatacctaac ctagtgatgg ttgatgatca tggtaaactt  480
gtgaatgaga atggtgttgg ggtcatacga agctatggag ctgatgctta tcctttcaca  540
ccagagaaaa tgaaggagat caaagaggat gaagatagag ctcggagaga acagacctta  600
agatctgtct tggtgactcc ttcacgagac tttgtgattt cgcctgacgg aaacaaggta  660
cccgtatcag agcttgaggg aaaaaccatt ggccttctct tctctgtggc ctcttacagg  720
aaatgcacag agcttactcc aaagcttgtt gagttttata cgaagctgaa ggagaataag  780
gaggattttg agattgtgtt gatatctctt gaagatgatg aggagtcttt taatcaagac  840
ttcaagacca agccatggct agcattgccg ttcaacgaca aaagtggatc aaaattggct  900
cggcatttca tgctgtcaac actaccgaca ctggtcattc tcggccctga tggaaaaatc  960
cgtcactcga atgtcgctga agctattgat gactatggag ttcttgcgta tcctttcact 1020
ccagagaagt ttcaagaact caaggagcta gaaaaggcaa aggtagaggc tcaaacgctc 1080
gagtcacttc ttgtctcagg tgatctcaac tacgttctcg gaaaagatgg ggcaaaggtg 1140
cttgtttcgg atctggtggg gaagactatt cttatgtact tctcagctca ctggtgtcct 1200
ccttgtcgcg cttttacacc aaagcttgtt gaagtataca agcagataaa ggagcggaat 1260
gaagcgtttg aattgatctt catctccagt gaccgtgacc aggaatcatt cgatgagtac 1320
tattcgcaaa tgccgtggct ggctcttcca tttggtgatc ctaggaaagc atccttggca 1380
aaaaccttta aggttggtgg atcccaatgc tagcagctct gggaccaact Gggcaaaccg 1440
taacaaaaga agcaagggac cttgtcgtag cccatggagc cgatgcttaC Ccccttttac 1500
tgaggaacgt ttgaaggaga ttgaagctaa gtacgatgag atagcaaaag attggcctaa 1560
gaaggtgaaa catgttctcc atgaagaaca tgagctagaa ctaactcgtg ttcaggttta 1620
cacatgcgat aagtgtgagg aagaagggac aatatggtga ccattgcgac gaatgcgact 1680
ttgatcttca cgccaagtgt gctttaaacg agtacacaaa agaaaacggc gatgaggctg 1740
tgaaagttgg tggcAgacga gtccaaagat ggttgggttt gtgaaggaaa cgtatgcacc 1800
aaggcctgat aggtagcttc aggcacagtt atgtcactat gtgtatgcaa agctatttcc 1860
tttctttgta taattgctaa tagtgagaat gtgtggttac aatctgatat ctgattaaga 1920
ccttggagat aaactctgtt aatctaaagt agtaataaaa gaatcttctt attt        1974
```

<210> 28406
<211> 470
<212> PRT